

Atty. Docket No. YOR920000168US1  
(590.014)

**REMARKS**

Applicants and the undersigned are most grateful for the time and effort accorded the instant application by the Examiner. The Office is respectfully requested to reconsider the rejections presented in the outstanding Office Action in light of the following remarks.

The disclosure continues to be objected to because of a number of asserted informalities. Equation 1 on Page 8 has been amended to correct a minor typographical error. The status identifier for Claim 2 has also been rewritten as "Previously Presented". Thus, it is submitted these objections should be withdrawn.

Claims 13 and 26 stand rejected under 35 USC 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. In particular, the Office is objecting to the variable C not being defined in the claim. Applicants respectfully traverse this rejection and assert the claims are not indefinite as C is understood by one of ordinary skill in the art. Thus, it is submitted this rejection should be withdrawn.

Claims 1-3, 6-12, 14-16, 19-25 and 27 stand rejected under 35 USC 103(a) over Goldenthal et al. in view of Newman et al. Specifically the Office asserted that "it would have obvious to one of ordinary skill in the art at the time the invention was made to modify Goldenthal by specifically providing multiple phonetic detail levels and the corresponding process, as taught by Newman, for the purpose of increasing efficiency and

Atty. Docket No. YOR920000168US1  
(590.014)

quality of a recognition system.” (*Office Action* at 6) Reconsideration and withdrawal of the present rejection is hereby respectfully requested.

The present invention broadly contemplates, in accordance with at least one presently preferred embodiment, the calculation of scores in such a way that the total likelihood is a weighted sum of the likelihood of all phonetic units at all levels of phonetic granularity (model grains), and that the weights are derived in such a way that the determination of the robustness and significance of the individual model grains is approached with emphasis. (Page 2, line 16 - Page 3, line 4) Given a structured model  $M(i,j)$  for a speaker with  $1 \leq i \leq L$  levels of detail and with  $1 \leq j \leq K(i)$  units on the  $i$ -th level, the score (as log-probability) for the utterance is calculated in each level separately, whereby explicit labeling information is used to identify the corresponding phonetic unit that is to be used on each level. (Page 8, lines 5-8) As discussed in the specification, the number of units on each level and the number of levels may vary across speakers, since there might be less data available from certain speakers, entailing the necessity of omitting certain units altogether. (Page 10, lines 7-9)

The previous comments regarding Goldenthal are equally applicable here. As presently best understood, Goldenthal appears to be directed to a two-stage cohort selection technique used to reduce the equal error rate of a speaker verification process which validates the claimed identity of an unknown speaker. (Col. 2, lines 61-65) First the digitized signals of an unknown speaker seeking verification are compared with acoustic models corresponding to the claimed identification to determine “claimed” log likelihood scores. (Col. 5, lines 18-22) Then the same testing signals are compared with

Atty. Docket No. YOR920000168US1  
(590.014)

all of the cohort models to determine cohort log likelihood scores and then a smaller subset of cohort scores is dynamically selected. (Col. 5, lines 23-30) The claimed scores and the dynamically selected scores are then presented to a validator which determines whether or not a threshold difference between the two scores is present. (Col. 5, lines 31-46) There is, however, no teaching or suggestion that the models used in calculating the various log likelihood scores have multiple levels of detail or that scores are calculated for each level of detail. In fact, the Office admits "Goldenthal fails to expressly disclose 'capable of having a plurality of phonetic detail of varying resolution for each frame' and determining a likelihood value for 'each level' of the phonetic detail of target speaker model." (*Office Action* at 6)

The previous comments regarding Newman are also equally applicable here. As presently best understood, Newman appears to be directed towards producing a speech model for use in determining whether a speaker associated with the speech model produced an unidentified speech sample. The speech model is produced without using an external mechanism to monitor the accuracy with which the contents were identified. (Col. 1, line 65 - Col. 2, line 11) As noted in the Office Action, each model of a word may be represented by a set of phonemes that represent the phonetic spelling of a word. (emphasis added) Furthermore, each phoneme may be represented by three sets of model parameters that correspond to the three nodes of the phoneme. (Col. 6, lines 26-29)

The Office Action, however, takes the position that this suggests "that the system includes multiple levels of phonetic detail and the corresponding processing for each level, as claimed." (*Office Action* at 6) Contrary to the position taken in the Office

Atty. Docket No. YOR920000168US1  
(590.014)

Action, however, there is no indication in Newman that each of the three sets of model parameters provides anything other than the same level of detail for each node of the phoneme. This is not, however, having multiple levels of phonetic detail in accordance with the present invention.

The present invention is “capable of having a plurality of levels of phonetic detail of varying resolution for each frame”. (Claim 1; emphasis added) However, as stated above, Newman appears to teach having multiple levels of phonetic detail for each word. Additionally, Newman appears to teach that each frame can be related to a particular phoneme node after disclosing that each word is represented by a series of phonemes. (Newman: column 3, lines 51-62). In the context of the instant invention, Newman appears to teach that each frame relates to one level of phonetic detail.

A 35 U.S.C. 103(a) rejection requires that the combined cited references provide both the motivation to combine the references and an expectation of success. There is, however, absolutely no teaching or suggestion in Newman that would lead one of ordinary skill in the art to modify Goldenthal to arrive at the present invention. Moreover, actually combining the teachings of Goldenthal and Newman would not result in the in the present invention which requires specifically “providing a model corresponding to a target speaker, the model being resolved into at least one frame and capable of having a plurality of levels of phonetic detail of varying resolution for each frame” and “determining, for each frame and each level of phonetic detail of the target speaker model, a likelihood value; and resolving the at least one likelihood value to obtain a likelihood score.” (Claim 1; emphasis added) Similar language appears in the

Atty. Docket No. YOR920000168US1  
(590.014)

other independent claims. This hierarchical approach in which there are a plurality of levels of phonetic detail with varying resolution, determining a likelihood value for each level, and then using the likelihood values to determine a likelihood score is simply not taught or suggested by either Goldenthal or Newman or the combination of the two.

Applicants acknowledge that Claims 4-5, 13, 17-18 and 26 were indicated by the Examiner as being allowable if rewritten in independent form. Applicants reserve the right to file new claims of such scope at a later date that would still, at that point, presumably be allowable.

In view of the foregoing, it is respectfully submitted that Claims 1, 14 and 27 fully distinguish over the applied art and is thus are in condition for allowance. By virtue of dependence from what are believed to be allowable independent Claims 1 and 14, it is respectfully submitted that Claims 2-13 and 15-26 are also presently allowable.

/

/

/

/

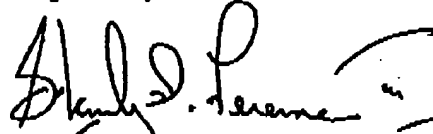
/

/

Atty. Docket No. YOR920000168US1  
(590.014)

In summary, it is respectfully submitted that the instant application, including Claims 1-27, is presently in condition for allowance. Notice to the effect is hereby earnestly solicited. In the unlikely event, however, it appears the claims will not be allowed, the Office is invited to call the undersigned to discuss the claims prior to the issuance of a further Office Action.

Respectfully submitted,



Stanley D. Ference III  
Registration No. 33,879

Customer No. 35195  
FERENCE & ASSOCIATES  
409 Broad Street  
Pittsburgh, Pennsylvania 15143  
(412) 741-8400  
(412) 741-9292 - Facsimile

Attorneys for Applicants